2020

A Study of Kindergarten Language Screening Performance Across Schools with Variable Demographic Compositions

Baylee Lackey
Western Kentucky University, baylee.lackey746@topper.wku.edu

A STUDY OF KINDERGARTEN LANGUAGE SCREENING PERFORMANCE ACROSS SCHOOLS WITH VARIABLE DEMOGRAPHIC COMPOSITIONS

A Capstone Experience/Thesis Project Presented in Partial Fulfillment

of the Requirements for the Degree Bachelor of Science

with Mahurin Honors College Graduate Distinction

at Western Kentucky University

By

Baylee J. Lackey

May 2020

*****

CE/T Committee:

Dr. Brian K. Weiler, Advisor

Dr. Allison Hatcher

Dr. O. E. Mansour

ABSTRACT

The Quick Interactive Language Screener (QUILS) for language impairment was designed and normed for children of diverse backgrounds. This study examined the utility of the QUILS in a diverse school district by comparing failure rates across five elementary schools. Schools varied in racial composition (non-White range: 20.45 – 80%) and socioeconomic status (free/reduced lunch qualified range: 35.35 – 100%). Among 321 currently-enrolled kindergartners, 272 completed the QUILS. Using author recommendations for kindergarten-aged cutoff scores, the district-wide screening failure rate for primarily monolingual English speakers (n = 212; via parent questionnaire) was 16.51% (range: 7.69 – 34.29% per school). Failure rates were not independent of school ($\chi 2(1, N = 5) = 16.92$, $p < .01$). Individual school QUILS failure rates significantly correlated with the percent non-White student population of the school ($r = .94$; $p < .01$) and the percentage of the school that qualified for free/reduced lunch ($r = .84$; $p < .05$).

Keywords: Quick Interactive Language Screener (QUILS), language development, kindergarten, demographics, language screening

To my wonderful friends, thank you for your admirable support
in times where I didn't even know I needed it.

To the faculty at Western Kentucky University, especially Dr. Weiler, thank you
for giving me the path to making my dreams a reality.

I dedicate this thesis to my family, who has shown me how to persevere through adversity and make efforts to always be kind. Thank you for always believing in me and pushing me to be the best version of myself.

ACKNOWLEDGEMENTS

I would like to extend my deepest thanks to the children and their families who elected to participate in this research. Additionally, I am grateful to the schools of the Bowling Green Independent School District for their support in data collection. The dedication I witnessed to children's futures from all platforms is truly admirable, and served to inspire me in completing this project. Thank you to Western Kentucky University for giving me an academic environment that quickly became a home on a hill. Thank you to the Mahurin Honors College for providing me the opportunity to make this thesis project a reality, and pushing me to elevate my undergraduate experience. Thank you to my thesis committee: Dr. Brian Weiler, Dr. Allison Hatcher, and Dr. O.E. Mansour for dedicating your time and efforts to my project.

I would like to extend gratitude to the faculty of the Communication Sciences and Disorders department for giving me the skills for success in a multitude of facets. Thank you to Dr. Brian Weiler for his unrelenting support throughout this process, and being an excellent role model for me as a future professional. I want to thank Katherine Church and Allyson Decker for being both my research partners and my friends in this project. Lastly, I would like to thank my friends and family for inspiring me in your unwavering support of my dreams. Thank you for showing me unconditional love and kindness in my life.

The data analyzed for this project came from research support to Brian Weiler provided by a Kentucky Biomedical Research Infastructure Network (KBRIN) grant funded by the National Institute of General Medical Sciences (NIH-NIGMS; 8P20GM103436).

VITA

EDUCATION

Western Kentucky University, Bowling Green, KY May 2020

B.S. Communication Disorders – Mahurin Honors College Graduate

Honors Capstone: *A Study of Kindergarten Language Screening Performance Across Schools with Variable Demographic Compositions*

Clarksville High School, Clarksville, TN May 2016

FIELDS OF STUDY

Major Field: Communication Sciences and Disorders
Minor Field: American Sign Language Studies
Master's Field of Study: Speech-Language Pathology

FIELD/WORK EXPERIENCE

WKU Internship Placement: Communication Disorders Clinic at Health Services
September 2019 to December 2019
Supervisor: Tammy Dockery, M.S. CCC-SLP

WKU Internship Placement: Kelly Autism Program
February 2020 to May 2020
Supervisor: Caroline Hudson, M.S., CCC-SLP

Boys and Girls Club of Bowling Green, Kentucky: Youth Development Professional
May 2018 to Present
Supervisor: Liz Bernard, CEO

WKU Department of Communication Sciences and Disorders: Research Assistant
July 2019- Present
Supervisor: Brian Weiler, Ph.D, CCC-SLP

PROFESSIONAL AFFILIATIONS

National Student Speech Language Hearing Association 2018- Present
Kentucky Speech-Language Hearing Association (KSHA) 2018- Present

INTERNATIONAL EXPERIENCE

WKU Study Abroad: Central Europe May 2017-July 2017
Poland, Austria, Czech Republic, Slovakia, Hungary

PRESENTATIONS

Lackey, B., Weiler, B., Decker, A., and Church, K. (2020, February). *A Study of Kindergarten Language Screening Performance Across Schools with Variable Demographic Compositions*. Poster presented at the 2020 Kentucky Speech-Language-Hearing Association Conference. Lexington, KY.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Current Status of Kindergarten Screening and Language Impairment Identification

As part of the 2004 reauthorization of the Individuals with Disabilities Education Act (IDEA), which formally recognizes the Response to Intervention (RTI) process, many schools utilize universally-administered grade-level screening tools to measure the academic readiness skills of kindergarten children (No Child Left Behind Act [NCLB], 2006). Falling under Tier 1 of the RTI process, kindergarten screeners for early literacy skills and math skills are now commonplace and have received attention from the research community for validation (e.g., Sittner Bridges & Catts, 2011; Brendefur et al., 2018). Failure of these screenings may indicate any number of phenomena, but in almost all cases is used in determination of further action by the caregiver or academic professionals. Further steps after failing a screener fall along a spectrum design, from progress monitoring to direct intervention (Weiler et al., 2018).

To date, the focus of universal screening at the kindergarten level has been on performance in reading and math rather than skills in speech and language. Lack of kindergarten readiness screening in language performance, however, may result in under-identification of critical oral language deficiencies that can adversely affect communication and academic outcome. In order for children at risk for language impairment (LI) to receive the services and supports required to minimize these adverse effects, they first need to be identified. Unfortunately, that is not often the case. In a

classic epidemiological study, less than one-third of kindergarten students with specific language impairment (or SLI) had been previously identified (Tomblin et al., 1997). Even further, only 9% of kindergarten children diagnosed with solely SLI and no associated speech sound disorders were reported by parents as having received speech-language services (Zhang & Tomblin, 2000). The identification of kindergarten-aged children with LI issue has not seemingly improved since Tomblin et al.'s study. (45% and 25%, respectively; Bishop & McDonald, 2009; Oetting, McDonald, Seidel, & Hegarty, 2016).

The current status of LI identification has resulted in recent advocacy for universal kindergarten language screening so that children with impaired language are not overlooked only to face adversity later in development (Rice, 2020). Identification of children at risk for LI is critical in light of the challenges faced by this population, including potential for poor literacy skills, reduced educational performance, social challenges, and future unemployment (Adlof & Hogan, 2019; Conti-Ramsden et al., 2013; Snowling et al., 2006; Whitehouse et al., 2009). Given these high stakes consequences, there is presently a call for research studies that contribute to the development and validation of measures appropriate for universal screening of oral language skills in school children from different backgrounds (Adlof & Hogan, 2019; Redmond et al., 2019).

Screening for LI carries benefits, but simultaneously poses a challenge in utilizing appropriate measures. One challenge facing administration is the applications involved with use in diverse populations. The 2014 U.S. Census reports a significant shift in the United States' demographic composition with the country undergoing a child majority-minority crossover in 2020. This projection additionally indicates that there will be no

racial or ethnic group with greater than 50% share of the nation's total by 2044, making the United States a plurality nation (Colby & Ortman, 2017). Some language assessment tools used in the identification process are not valid for use in culturally and linguistically diverse populations (Weiler et al., 2018; Qi et al., 2006). Studies of cultural- and linguistic-minority children (e.g., child speakers of African American English dialect; AAE) reveal unmanageably high language screening failure rates of around 50-60% or greater with conventional assessments and scoring methods (Craig & Washington, 2004; Hendricks & Adlof, 2017; Moland, 2011). Such elevated failure rates suggest that the screening assessments were not sensitive to AAE dialect features (which are not inherently indicative of impairment) and therefore resulting in an inordinate number of false positives. Instead, although it is recommended that screeners over-identify rather than under-identify in order to minimize false negatives,(Klee, Pearce, & Carson, 2000), the suggested universal kindergarten language screening failure rate is < 30% given the 7-10% prevalence of childhood LI (Oetting, Gregory, & Rivière, 2016). This < 30% percentage is based on Tomblin et al.'s (1997) screening failure rate of 26.2% (7,218 kindergarteners screened).

The Quick Interactive Language Screener (QUILS; Golinkoff, de Villiers et al., 2017) is unique in that its development incorporated an intended use in linguistically diverse settings. The QUILS is a research-based language screener for children ages 3-5 that examines a child's comprehension skills across three language domains:

1. Vocabulary (words children understand; 16 items)
2. Syntax (knowledge of the way words are put together in sentences; 16 items)
3. Process (skill in rapidly learning new language information; 16 items)

These areas of language are critical in communication, as they describe how we use and learn language. The developers of the QUILS were inspired by the need for an efficient (i.e., 10-15 minute) evidence-based screening tool to measure language in preschool and kindergarten children (Golinkoff et al., 2017). The QUILS utilizes standardized norms that can be used to flag children who may be falling behind in critical language areas as compared to their peers.

As described in the QUILS user's manual, this research-based screening is designed to accommodate for linguistically, socioeconomically (SES), and culturally diverse populations. The user's manual further reports that, "the items of the QUILS were selected through careful testing to be culturally and dialectally neutral; they do not place children from a range of cultural backgrounds and children who speak dialects such as African American or Appalachian English at a disadvantage" (Golinkoff et al., 2017, pp.16). Accordingly, the test item words and linguistic structures on the QUILS were carefully and methodically chosen to be culturally- and dialectally-neutral so as not to be biased against African-American English (AAE) or Spanish-influenced English. For example, even though difficulty with past tense *–ed* is a clinical marker for LI in MAE-speaking children (e.g., Rice, Wexler, & Hershberger, 1998), the rules of AAE do not always require past tense *–ed* (Pruitt & Oetting, 2009). Accordingly, so as not to potentially bias against child speakers of AAE (i.e., confuse dialect with disorder), the QUILS only assesses the past tense copula and auxiliary verb "*was*" (e.g., "Where *was* the hat" instead of "What *happened* to the hat?") since it is obligatory in both MAE and AAE (see Figure 1).

Figure 1. Dialectically neutral QUILS test item from the Syntax domain

Past Tense Item 7. "Look at this hat. Uh oh, it blew off! Where was the hat?"

Additionally, children from low SES and/or minority dialect backgrounds may have low existing vocabulary knowledge (due in part to limited home language and literacy exposure) yet nonetheless have age-appropriate abilities in learning new words and structures. In order to reduce potential cultural-linguistic bias involved in testing only existing language knowledge, the QUILS examines children's ability to conduct processing operations that are minimally dependent on prior knowledge or experience so that all children, irrespective of background, are on comparable footing (Campbell et al., 1997). For example, on the noun learning Process items of the QUILS, children must *fast map* a novel label (e.g. *gelp*) onto an object and then *extend* that label to another member of the same category (see Figure 2). Difficulties with the process of fast mapping are characteristic of children with LI but not necessarily of children from cultural/linguistic minority groups. For example, Rice et al. (1994) found that, after ten exposures to a new word, the 5-year-old LI group retained comprehension of less than half the words as the 5-year-old typical language (TL) group ($p < .001$), performing more like the 3-year-old TL group. By contrast, in a study of fast mapping among African American preschoolers, Horton-Ikard & Ellis-Weismer (2007) reported no difference between low-SES and

middle-SES groups on novel word learning despite a significant difference favoring the middle-SES group on measures of existing vocabulary knowledge (*p*s = .004, .001). Therefore, the QUILS has potential for valid use as a universal screener for kindergarten language impairment in diverse populations.

Figure 2. Noun learning item from the Process domain of the QUILS

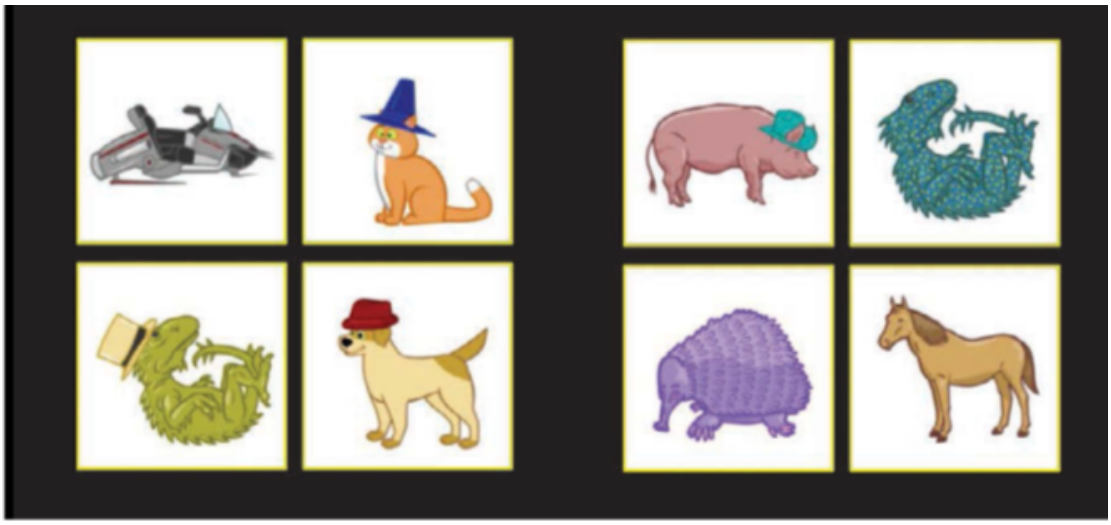


Figure 4.11. Noun Learning Item 35. Part 1 (Fast mapping): "A gelp is wearing a hat. Show me the gelp with the hat." Part 2 (Extension): "Ca show me another gelp?"

The racial composition of the QUILS normative sample was designed to be representative of monolingual English-speaking children in this age range in the U.S. (Table 1). In addition to the monolingual sample, 23.3% of the children from the normative sample were reported by parents to be of Hispanic origin. The percentage of mid-socioeconomic status (SES) in the QUILS normative sample is comparable to 2014 U.S. census data for maternal aged females (i.e., 18-39 years) with at least an associate's degree.

Table 1

Composition of the monolingual norming sample for the QUILS

QUILS Norming Sample	
Total Children	415
Race/Ethnicity	
African American	31.6%
White	57.8%
Multiracial	8.8%
Asian	<1%
Other	<1%
Socioeconomic Status (SES)	
Low: *n* (%)	254 (61.20)
Mid: *n* (%)	160 (38.55)
High: *n* (%)	1 (.24)

A challenge facing screening procedures is the determination of what cut point (i.e., criterion score) will be used to operationalize failure. It is commonplace to use deviation below an expected mean performance as guidance in making this determination. However, in a study on 33 currently-used language assessments, researchers found that applying a conventional cutoff score to interpretation of assessment results (typically -1.0 to -1.5 standard deviations below the mean) is unsupported by the evidence available in manuals (Spaulding, Plante, & Farinella, 2006).

The researchers who developed the QUILS created the screener's cut off scores for failure based off the performance of the normative sample. The QUILS was normed on 415 children (216 male) aged 3;0 to 5;11 (M = 4;5; *SD* = 0;9). According to the QUILS manual, failure is defined as a percentile score below the 25th percentile of the normative sample at each one year chronological age interval. Since the overarching purpose of a screener is to identify risk and not to definitively diagnose, the QUILS development team delineated a 25th percentile cut score to be a conservative estimate of

risk in light of language impairment prevalence estimates of approximately 7-10% (Leonard, 2014; Rice et al., 2020). QUILS software automatically converts a child's raw score (i.e., the total number correct out of 48 items) to an age-based (e.g., 5;0-5;11) standard score and corresponding percentile rank. Percentile ranks (and standard scores) are automatically generated for the child's overall performance (i.e., all 48 items) as well as for each of the three areas assessed (Vocabulary, 16 items; Syntax, 16 items; Process, 16 items). Per manual recommendations, QUILS failure is obtained by performance below the 25th percentile on: (a) the overall score (b) both the Vocabulary and Syntax areas, or (c) the Process area.

The Current Investigation

The QUILS carries potential for use as a universal language screener for diverse kindergarten children, who are typically 5 years of age at the point of school entry. The purpose of the broader study that provided data for the present investigation is to examine the reliability and validity of the QUILS when used as a universal kindergarten language screener in a diverse school district. The racial and SES makeup of the small city public school district, Bowling Green Independent School District (BGISD; Table 2), that participated in the kindergarten language screening study is comparable to that of QUILS normative (Table 2). Specifically, the relative overall percentages of White and Black children are comparable across the two populations. Additionally, both populations are comprised of a majority of children from low-SES households, shown in Table 2 as the percentage of children from each of the five BGISD elementary schools eligible for free or reduced lunch (National Center for Education Statistics, 2019).

Table 2

Demographics of Bowling Green Independent School District

	Bowling Green Independent School District (BGISD)					
Elementary School		**School A**	**School B**	**School C**	**School D**	**School E**
Total Students	4,206	396	281	445	335	513
Kindergarten Students (2017-18)		69	39	79	52	84
Race/Ethnicity						
African American	18.95%	14.39%	16.01%	6.74%	38.51%	23.98%
Hispanic	16.88%	5.05%	11.39%	3.37%	29.55%	32.75%
White/Caucasian	53.54%	67.17%	60.14%	79.55%	20.00%	29.82%
Native American	0.38%	0.51%	0.36%	0.22%	0.30%	1.36%
Asian/Pacific Islander	5.25%	3.79%	3.20%	4.94%	4.18%	7.02%
Two or More Races	4.99%	9.09%	8.90%	5.17%	7.46%	5.07%
Free Lunch Eligible		30.3%	75.09%	24.72%	99.10%	92.40%
Reduced Lunch Eligible		5.05%	6.76%	3.15%	2.39%	2.53%

There are several components that contributed to the development of this present study. One such component is the selection of the normative sample by which to evaluate child performance on the QUILS. As described above, scoring (i.e., pass/fail) recommendations at the 25th percentile relative to QUILS normative sample reference database are provided by the test authors. The authors reinforce, however, that scoring is not concrete and that individual schools or school districts may elect to use different approaches. One such approach is the development of cutoff scores at the local level by examining the distribution of scores to create local norms. Local norms, as compared to published test norms, can be advantageous in that they are more representative of the sample being tested (e.g. Tomblin, Records, & Zhang, 1996). Using local norms provides clinically useful information that adjusts standard norms for cultural, linguistic, or social-based fluctuations, making it relevant in applications to this study and the QUILS's accuracy (Plante & Vance, 1995).

An additional component considered in the development of the present investigation is the manner in which the QUILS screening failure outcomes align with existing target failure rates of kindergarten screening. In the previously mentioned study, Tomblin et al. (1997) places the prevalence of LI in kindergarten populations between 7% and 12% (7,218 kindergarteners screened). Even further, it is understood that when giving a screening it is preferable to over-identify rather than under-identify in order to minimize false negatives, (Klee, Pearce, & Carson, 2000). The suggested universal kindergarten language screening failure rate is $\leq$ 30% given the 7-10% prevalence of childhood language impairment (Oetting, Gregory, & Rivière, 2016). In sum, it is better

to attempt at identifying all those students at risk of LI rather than miss vulnerable children who may miss diagnosis due to absence of a reliable screening.

The QUILS has potential to serve as a universal kindergarten language screener in an environment of diversity. To better understand this potential, research was conducted within the Bowling Green Independent school district in the area of kindergarten language screening using the QUILS. Preliminary data was collected at five different schools that vary in demographic composition. Based on reported statistics, these schools evidence notable contrasts in both racial composition and socioeconomically disadvantaged student enrollment. For example, School D is comprised of roughly 80% non-White minority students and nearly 100% of students qualify for free/reduced lunch. In comparison, School C in the same school district is reported as being comprised of 20.45% non-White minority students and 27.87% of students qualify for free/reduced lunch (Table 2, National Center for Education Statistics, 2019). This range will allow for the examination of a possible correlation between QUILS failure rates and the demographic makeup of individual schools. Such an examination is important in establishing the validity of the QUILS as a non-biased kindergarten language screening tool across schools of varying demographic makeups.

The purpose of the current study is to evaluate the kindergarten language screening failure rates in a demographically diverse school district using a measure (QUILS) that was specifically designed to minimize cultural, socioeconomic and dialectical bias. The primary research questions posed in this study are:

1. Do schools with greater minority populations and/or increased qualification for free/reduced lunch evidence increased rates of failure on the QUILS?

2. What is the percent failure rate at the 25th percentile for each school using both QUILS published norms and local norms?

Expected Results and Significance

One expected outcome of this research is to provide evidence of a relationship between demographic makeup of the school and failure rates on the QUILS. The QUILS was designed as a minimally-biased method of screening language in culturally and economically diverse populations and is predicted to be valid measure of risk for language impairment in children regardless of cultural background. Given this design, we would not expect there to be a significant discrepancy between schools with increased populations of free/reduced lunch qualified and non-white minority students and those with decreased free/reduced lunch qualified and non-white student makeups (e.g., School D as compared to School C). In turn, this would potentially support an updated approach to utilizing linguistically unbiased language screening tools for in the field. On the other hand, the analyses may yield a correlation between failure rates and school racial/SES diversity. Prior research points to such a correlation for the vocabulary and syntax domains of language (e.g., Qi et al. 2006; Qi et al., 2003), but not necessarily for the process domain of language (Horton-Ikard & Ellis-Weismer, 2007). If a significant positive correlation between QUILS failure rates and low school SES and/or QUILS failure rates and high school minority population is found, this project would help to clarify the potential impact of increased diversity on language impairment screening failure. These results could progress research in speech-language pathology for culturally diverse populations with a variety of dialects by demonstrating a need for building language skills in increasingly diverse schools.

Regardless of the outcome, this study will contribute to an understanding of universal screening for detection of risk for language impairment in demographically diverse populations. As previously mentioned, the field of speech-language pathology is shifting to a new view on diverse clientele which includes an examination of dialects within communication disorders rather than separating the two (Oetting, Gregory, & Rivière, 2016). These dialects are present in current kindergarten populations, creating an additional barrier to the current existence of unidentified language impairments in kindergarten. This research is imperative due to the academic needs of elementary-aged children with language impairment. For these children to benefits from intervention services from speech-language pathologists and educators, they must first be identified.

CHAPTER 2

METHOD

This project was approved by the Institutional Review Board at Western Kentucky University (#1515751-1).

Procedure

Data from an existing database on QUILS performance (Weiler, 2019; WKU IRB# 983757-3) was examined for pass/fail rates by elementary school site. This data was collected by a team of students and Dr. Weiler at five separate elementary schools within the Bowling Green Independent school district as part of a broader study on detection of child language impairment. All BGISD kindergarten families at the start of the 2019-2020 school year were invited to have their child participate. Of the 321 kindergartners enrolled at the beginning of the school year, 84.42% (n = 271) completed the QUILS (Range: 78.26% - 95.93% participation per school). The valid use of local norms is dependent upon sample size distributions of at least 100 participants for study adequacy (McCauley & Swisher, 1984), thus validating the sample size for this study.

The QUILS was individually-administered in the English language via an automated program loaded on a touchscreen tablet (iPad) at each respective school. Audio stimuli for each of the 48 items of the QUILS were presented to the child through child-sized on-the-ear headphones (SONY MDR-222KD). A member of the research team wore a pair of headphones as well that were connected to the iPad by an audio splitter to ensure that the QUILS audio was functional at all points during the

administration. When necessary, for example when there was an announcement over the school's loudspeaker, the research team member was able to briefly pause the QUILS administration. Administration time for each child ranged from approximately 10-15 minutes.

For the purpose of this study, students classified as not primarily monolingual English speakers via parent report on the QUILS language questionnaire (see Appendix) were not included in the analysis since screening failure among this group may be the result of limited English language exposure. The QUILS language questionnaire was provided in English and, when appropriate, Spanish. Following guidelines from the QUILS manual, 59 kindergartners who were administered the QUILS were excluded from the present study because they were not primarily monolingual English speakers. Of the remaining 212 kindergartners whose QUILS performance was considered for the present study, the mean age was 5 years, 6 months (SD = 4 months) and 47.17% were girls. The mean ages and gender distributions of participating monolingual kindergartners at each BGISD elementary school are listed in Table 3.

Table 3

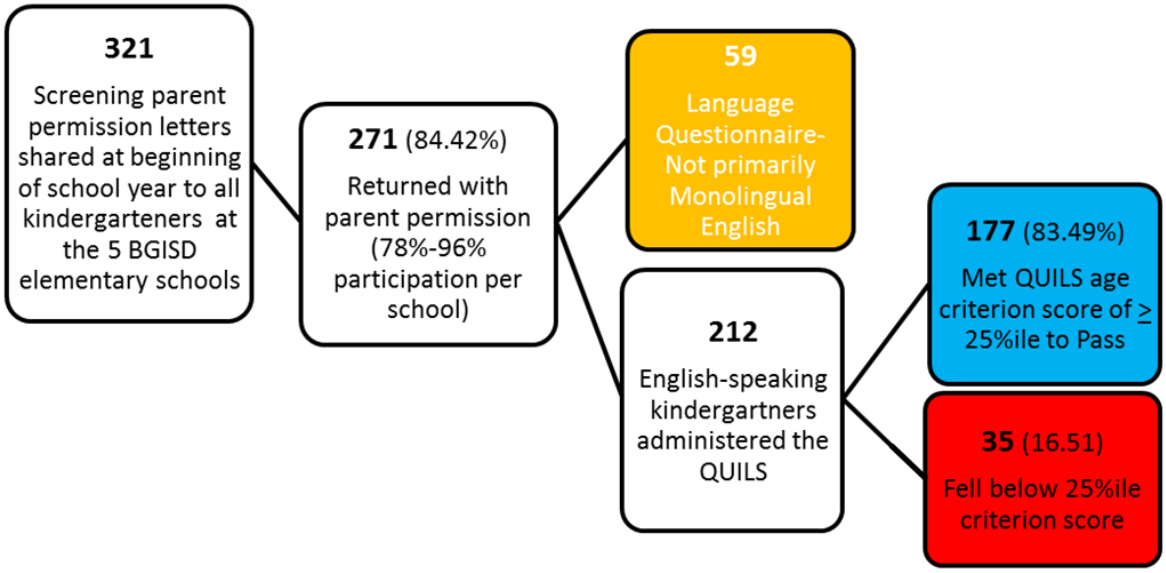
Number of Children Screened, Mean Age, and Gender Breakdown per School

	School A	School B	School C	School D	School E
Number of kindergartners screened (monolingual)	52	33	66	26	35
Mean Age (years; months)	5;6	5;6	5;7	5;6	5;7
(SD in months)	(4)	(4)	(5)	(4)	(3)
Percentage of girls screened	48%	52%	50%	39%	43%

Using QUILS manual recommendations for kindergarten-aged (i.e., 5;0 – 5;11) cutoff scores at the 25th percentile, the screening failure rate for primarily monolingual

English speakers was 16.51% (n = 35; Figure 3). The mean age (SD) of the children who failed (5;6 (SD = 4 months)) was comparable to that of the children who passed (5;7 (S = 4 months)). Of the children who failed, 37% were girls. Of the children who passed, 49% were girls. QUILS percentile scores were automatically generated by the QUILS program from raw scores that were saved for each participant based on his or her touchscreen responses to each of the 48 screener items. As a measure of reliability, a random 20% of the QUILS scores (n = 42) were additionally hand calculated by summing the stored individual item responses on the QUILS program and then using the raw score to standard score to percentile rank conversion charts provided in the manual. Results of this scoring reliability check showed 100% agreement between the automatically generated percentile ranks and those calculated by hand.

Figure 3. Participant flow chart from recruitment through completion

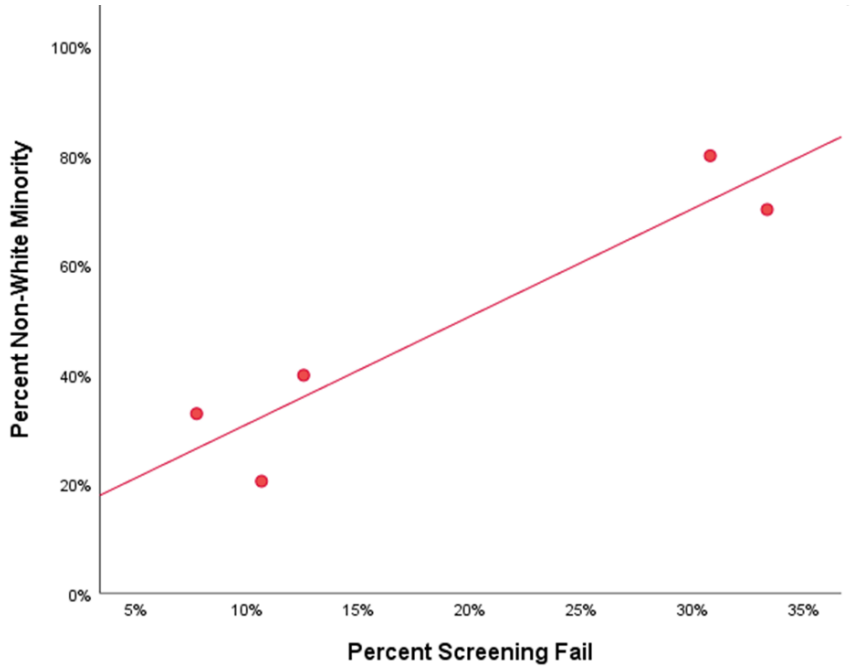


CHAPTER 3

RESULTS

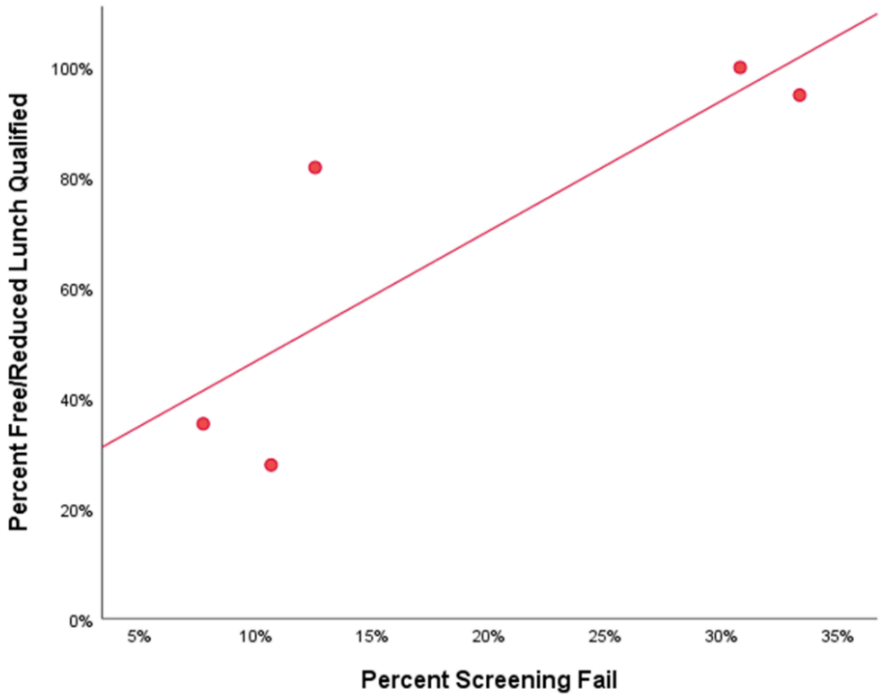
To address the first research question, a bivariate correlational analysis was carried out to determine the relationship between school non-White minority population and school kindergarten QUILS failure rates. For this question, failure was established using published QUILS norms and defined as performance below the 25th percentile on: (a) the overall score (b) both the Vocabulary and Syntax areas, or (c) the Process area. QUILS failure rates positively and significantly correlated with the percent non-White student population of the school (r = .94; p < .01). Schools with greater percentages of non-White students had greater QUILS failure rates (Figure 4).

Figure 4. Regression line showing the correlation between the percentage of screening failures per school and the percent non-White minority population per school.



To further address the first research question, a bivariate correlational analysis was carried out to determine the relationship between percent school qualification for free/reduced lunch and school kindergarten QUILS failure rates. QUILS failure rates positively and significantly correlated with the percentage of the school that qualified for free/reduced lunch ($r = .84$; $p < .05$). Schools with greater percentages students qualifying for free/reduced lunch had greater failure rates (Figure 5).

Figure 5. Regression line showing correlation between the percentage of screening failures per school and the percent of the school population that is free or reduced lunch qualified.



To address the second research question, descriptive statistics were used to quantify the percent of children at each school who failed the QUILS using bot published QUILS norms as well as local norms. As mentioned above, the total QUILS failure rate for the entire participant sample using published norms and the recommended 25th percentile cutoff was 16.51% (n = 35 fail; n = 177 pass). When considering failure rates

at the individual school level, it is not surprising to see variability given the strong correlations reported for the first research question. Indeed, the percentage of children that failed the QUILS by any means (i.e., by Overall score, by Process score alone, and/or by Vocabulary + Syntax scores combined) ranged from 7.69% - 34.29% (Table 4). Since the dependent variable for the second research question is categorical, that is, the outcome is either pass or fail, a Chi-squared test was run to determine whether the observed school failure rates were significantly different from the expected failure rates if QUILS failure rates are independent of school. The Chi-square test result shows that, in fact, QUILS failure rates (by any means) were not independent of school ($\chi 2(1, N = 5) = 16.92$, $p < .01$). As appreciated in Figure 6, the two schools with both very high poverty and high non-White minority populations had failure rates approximately three times greater than the other schools. Despite this range, individual school failure rates at all five elementary schools fell around or below the recommended language screening ceiling rate of ~30% (dashed red line, Figure 6) but above the 7% prevalence estimate for primary language impairment (i.e., SLI, DLD). This is particularly the case when looking at failure rates for the Process score alone, which fell between 7 – 30%. Recall that the Process domain of the QUILS assesses rapid word learning skills that are thought to be less impacted by a child's home language environment than, for example, known vocabulary. Therefore, it was expected that QUILS failure rates when considering exclusively Process score might be more comparable across schools varying by SES. That was not the case. A Chi-squared test of QUILS failure rates by Process score revealed that failure rates were not independent of school ($\chi 2(1, N = 5) = 9.96$, $p < .05$).

Table 4

QUILS Failure Rates using Published Norms by School and Failure Type

Failure Rates According to Standard Norms					
Elementary School	**School A**	**School B**	**School C**	**School D**	**School E**
Failure Type					
Overall Score	1.92%	9.09%	3.03%	19.23%	20.00%
Process Score	7.69%	12.12%	7.69%	23.08%	28.57%
Vocabulary/Syntax Scores	0.00%	3.03%	0.00%	7.69%	11.43%
By Any Means	7.69%	12.12%	10.61%	30.77%	34.29%

The second research question was also addressed using local norms. Given the adequate sample size of at least 100 participants in the data set, local norms were established following conventional procedures (Baumgartner, 2009; Chew, Kesler, & Sudduth, 1984). Again using the 25th percentile at the cutoff, failure rates for individual schools were calculated (Table 5). The total QUILS failure rate for the entire participant sample using local norms was 29.27% (n = 63 fail; n = 149 pass). The percentage of children that failed the QUILS by any means (i.e., by Overall score, by Process score alone, and/or by Vocabulary + Syntax scores combined) ranged from 21.21% - 53.85% (Table 5). For each school and the district overall, QUILS failures were more frequent using local norms as compared to when QUILS published norms were applied. A Chi-squared test was run to determine whether the observed school failure rates using local norms were significantly different from the expected failure rates if QUILS failure rates are independent of school. The Chi-square test result shows that QUILS failure rates (by any means) using local norms were not independent of school ($\chi 2(1, N = 5) = 16.06$, $p < .01$). As appreciated in Figure 7, the two schools with both very high poverty and high

non-White minority populations had failure rates approximately two times greater than the other schools. Additionally, when considering failure by any means, these two elementary schools fell markedly above the recommended language screening failure ceiling rate of ~30% (Oetting et al., 2016). Turning to the Process score alone, failure rates using local norms for the two very high poverty schools dropped, but still fell above 30%. A Chi-squared test of QUILS failure rates by Process score using local norms revealed that failure rates were not independent of school ($\chi 2(1, N = 5) = 10.02$, $p < .05$).

Table 5

QUILS Failure Rates using Local Norms by School and Failure Type

Failure According to Local Norms					
Elementary School	**School A**	**School B**	**School C**	**School D**	**School E**
Failure Type					
Overall Score	15.38%	21.21%	13.64%	50.00%	42.86%
Process Score	15.38%	18.18%	19.70%	34.62%	40.00%
Vocabulary/Syntax Scores	7.69%	15.15%	4.55%	34.62%	22.86%
By any Means	23.08%	21.21%	21.21%	53.85%	45.71%

Figure 6. Screening failure rates by school and failure type using QUILS published norms.

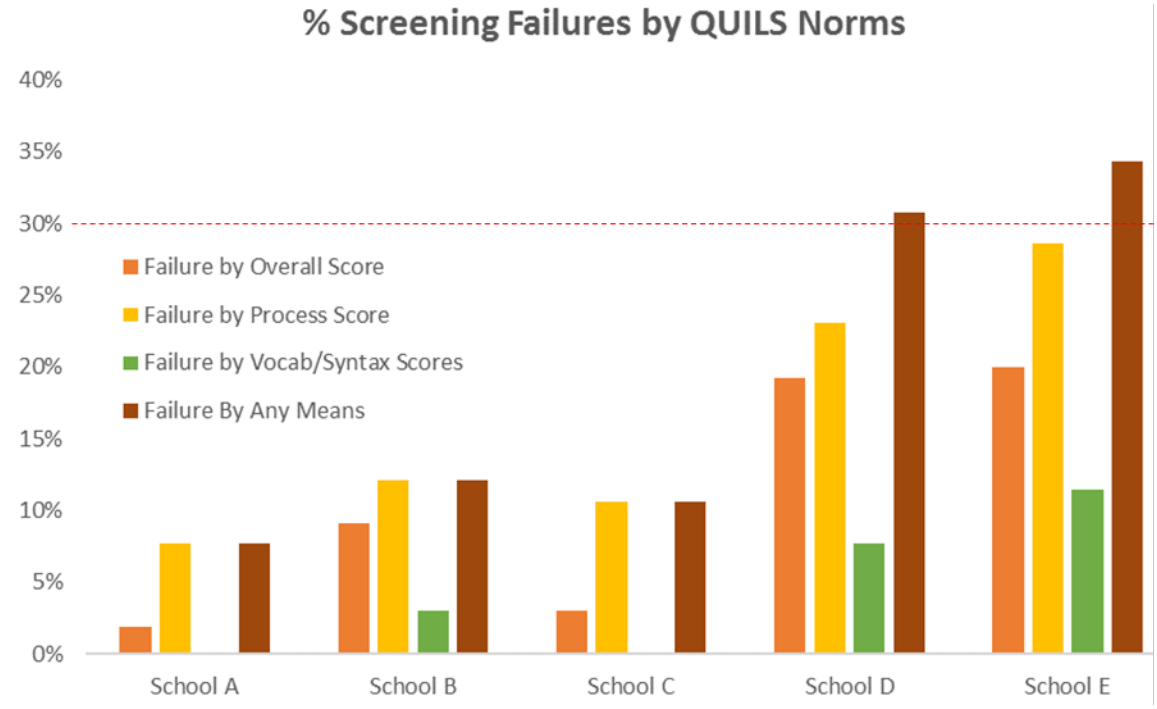
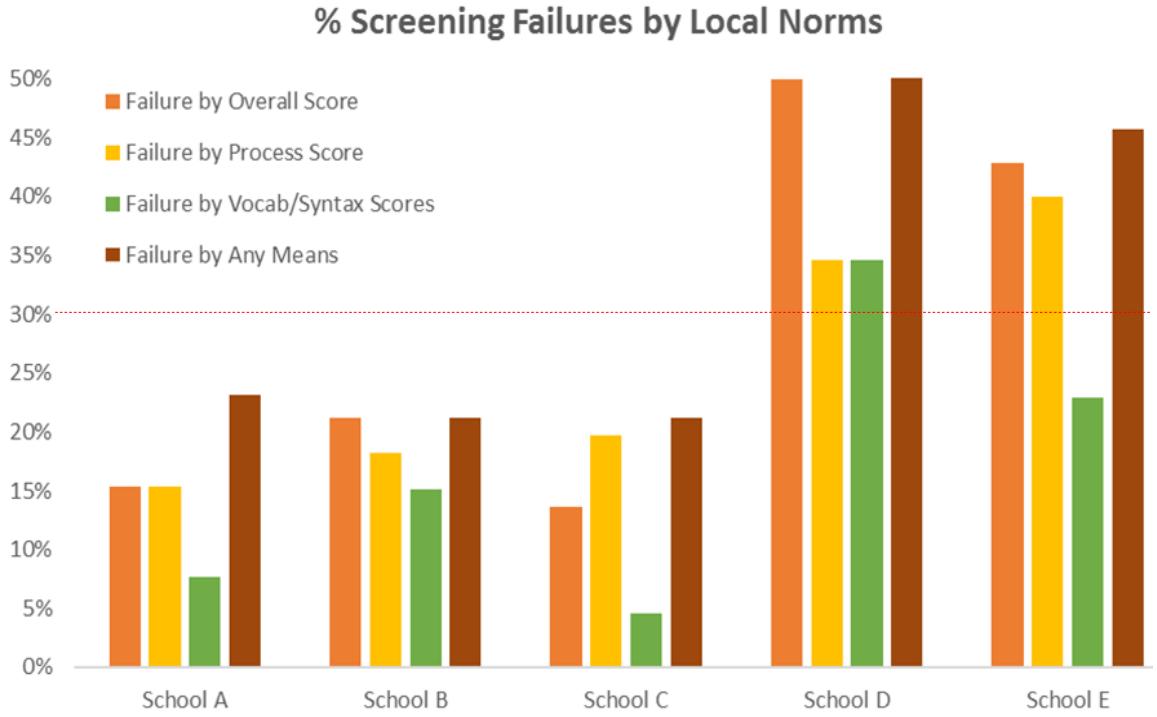


Figure 7. Screening failure rates by school and failure type using local norms.



CHAPTER 4

DISCUSSION

The purpose of this study was to examine whether schools with greater non-White minority populations and/or increased qualification for free/reduced lunch evidenced increased rates of failure on the QUILS. In doing so, the validity of the QUILS as a non-biased kindergarten language screening tool when used in schools with diverse demographic makeups was evaluated. The rationale for this research stems from a need for progression in our understanding of universal language screening techniques in diverse kindergarten populations. Driven by this motivation, the results of the present study revealed two key findings.

1. Significant correlations were found between screening failure percentages and the percent minority population of each school as well as the percent free/reduced lunch qualified of each school.
2. Using QUILS norms, individual failure rates at all five elementary schools fell around or below the recommended language screening ceiling rate of ~30% but above the 7% prevalence estimate for primary language impairment, providing partial evidence for the validity of the QUILS as a universal kindergarten screener in a diverse school district.

Correlations in Failure Percentages and Demographics

Given prior evidence of low SES disadvantages in vocabulary knowledge (e.g., Farkas & Beron, 2004; Taylor et al., 2013) and syntax skills (e.g., Huttenlocher et al.,

2010; Vasilyeva, Waterfall, & Huttenlocher, 2008), it was not unexpected to see higher QUILS failure rates in the schools with the highest percentages of children qualifying for free/reduced lunch. These schools (Schools D and E) also have the highest non-White minority populations, making it difficult and beyond the scope of the present study to tease out the unique effects of race and SES on QUILS failure rates. For children from mid-SES schools (i.e., those with lower free/reduced lunch qualified rates such as Schools A or C) having more robust prior experience and skills in vocabulary and syntax may have placed them at an advantage for screening pass (e.g., Qi et al. 2006; Qi et al., 2003). This possibility is consistent with results of both the correlational analyses of QUILS failure rates as well as Chi-squared analyses showing that failure rates (by any means) were not independent of the school.

However, the findings specific to the Process category of the QUILS were unpredicted upon initiation of the study. Based on prior research, the Process domain of language was not anticipated to show strong SES differences across schools (e.g., Burton & Watkins, 2007; Horton-Ikard & Weismer, 2007). However, screening failure rates by Process alone were also not independent of school. ($\chi 2(1, N = 5) = 9.96$, $p = .04$), with low SES schools again showing increased failure rates Despite the design of the QUILS to minimize racial and economic bias, results show that all three categories of the QUILS interplayed to disproportionately identify children from high poverty and high non-White schools..

The relationship between the demographic composition of the educational setting and the likelihood of failure on the QUILS is consistent with another recent report. Levine et al. (2018) utilized the QUILS to examine the effects of SES on vocabulary,

syntax, and process in three- to five-year old children. Results from Levine et al.'s study demonstrated that low SES status had a comparable negative impact on all three categories of language examined by the QUILS., implying that demographic differences have pivotal impact on kindergartener's performance in language skills. In sum, the positive correlations found between low SES/high school minority rates and QUILS failure rates expand the knowledge base of the potential impact of demographic factors on language on screening performance.

Validity of the QUILS

The findings of this study nonetheless provide evidence for the validity of the QUILS as a universal kindergarten language screener. Using published QUILS norms, individual failure rates at all five elementary schools fell around or below the recommended language screening ceiling rate of ~30% but above the 7% prevalence estimate for primary language impairment when QUILS norms were applied. However, when failure rates were examined using local norms, results yielded findings that did not fall in accordance with recommendations for prevalence and failure rate. In particular, failure rates (by any means) for low SES/high minority schools hovered around 50% which is unmanageably high for the purpose of a universal screening tool. As was the case with the utilization of QUILS published norms, the results of the Chi-squared analyses using local norms revealed that failure rates, both by any means and by Process alone, were not independent of school. Overall, the comparison of failure rates by published norms and failure rates by local norms suggests that the norms and recommendations provided by the QUILS manual are perhaps more applicable for use in populations with high diversity rates like BGISD. As noted in Table 1, whereas 61.25%

of the children from the QUILS norming sample were low SES households, only 44.34% of the 212 BGISD kindergartners in the current study were from low SES schools (Schools B, D, and E). Accordingly, the majority of the QUILS scores (55.66%) contributing to the local norms distribution came from higher performing mid-SES schools (Schools A and C). Given the strong correlation between QUILS failure rates and SES, it thus stands to reason that low SES school failure rates using BGISD local norms would exceed those using QUILS norms.

Limitations

There are several limitations to be considered in this study. One way in which this study is limited is by the use of cross-sectional correlational analysis rather than longitudinal designs. An obvious next step in this line of research would be an examination of how the factors involved (SES, race, language) develop over time. Additionally, this study did not address how the elements of language examined by the QUILS impact or relate to each other. Another limitation presented in this study is the absence of validity testing to verify the language status of children relative to their performance on the QUILS. Confirmatory validation testing using a reference standard test for language impairment (CELF-5) administered to a random subset of participants from the present study who both passed and failed the QUILS screener is currently being carried out as part of the broader study that sourced the data for the present investigation.

Clinical Implications and Future Directions

Completion of this project contributes to the knowledge base required in order for the QUILS to be considered for utilization for identifying children at risk for LI. Results of this work carry the potential to increase clinical understanding of the benefits and

limitations of using the QUILS in culturally diverse populations. This information may additionally contribute to the field of speech-language pathology by providing a basis for further study on the impacts of diversity on kindergarten screening outcomes. Replicating the study with longitudinal aspects is one aim for future research which has potential to yield significant results. By adding longitudinal factors such as samples from BGISD for children in 1st grade or administration of comprehensive language assessment, projected research may further contributions to the field.

Conclusion

This study provides evidence for both the outcomes of universal language screening in kindergarten populations as well as the relationship between school racial and economic diversity and QUILS' performance. Findings from this study carry implications for clinical practice in the field while simultaneously opening an opportunity for further research in the components of this project to strengthen professional practice. There remains a definite need for continual progress in the area language screening procedures applied to culturally diverse populations.

BIBLIOGRAPHY

Adlof, S. M., & Hogan, T. P. (2019). If We Don't Look, We Won't See: Measuring Language Development to Inform Literacy Instruction. *Policy Insights from the Behavioral and Brain Sciences*, *6*(2), 210-217.

Baumgartner, T. A. (2009). Tutorial: Calculating percentile rank and percentile norms using SPSS. *Measurement in Physical Education and Exercise Science*, *13*(4), 227-233.

Bishop, D. V., & McDonald, D. (2009). Identifying language impairment in children: combining language test scores with parental report. *International Journal of Language & Communication Disorders*, *44*(5), 600-615.

Brendefur, J. L., Johnson, E. S., Thiede, K. W., Strother, S., & Severson, H. H. (2018). Developing a multi-dimensional early elementary mathematics screener and diagnostic tool: the primary mathematics assessment. *Early childhood education journal*, *46*(2), 153-157.

Burton, V. J., & Watkins, R. V. (2007). Measuring word learning: Dynamic versus static assessment of kindergarten vocabulary. *Journal of Communication Disorders*, *40*(5), 335-356.

Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech, Language, and Hearing Research*, *40*(3), 519-525.

Chew, A. L., Kesler, E. B., & Sudduth, D. H. (1984). A practical example of how to establish local norms. *The Reading Teacher*, *38*(2), 160-163.

Colby, S. L., & Ortman, J. M. (2017). Projections of the size and composition of the US population: 2014 to 2060: Population estimates and projections.

Conti-Ramsden, G., Mok, P. L., Pickles, A., & Durkin, K. (2013). Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. *Research in developmental disabilities*, *34*(11), 4161-4169.

Craig, H. K., & Washington, J. A. (2004). Grade-related changes in the production of African American English. *Journal of Speech, Language, and Hearing Research*.

Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research*, *33*(3), 464-497.

Golinkoff, R. M., De Villiers, J. G., Hirsh-Pasek, K., Iglesias, A., Wilson, M. S., Morini, G., Brezack, N. (2017). *User's Manual for the Quick Interactive Language Screener (QUILS): A Measure of Vocabulary, Syntax, and Language Acquisition Skills in Young Children*. Paul H. Brookes Publishing Company.

Hendricks, A. E., & Adlof, S. M. (2017). Language assessment with children who speak nonmainstream dialects: Examining the effects of scoring modifications in norm-referenced assessment. *Language, Speech, and Hearing Services in Schools*, *48*(3), 168-182.

Horton-Ikard, R., & Weismer, S. E. (2007). A preliminary examination of vocabulary and word learning in African American toddlers from middle and low socioeconomic status homes. *American Journal of Speech-Language Pathology*.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive psychology*, *61*(4), 343-365.

Individuals with Disabilities Education Improvement Act of 2004. P.L. 114-95, 118 Stat. 2647 (2004 Dec 3).

Klee, T., Pearce, K., & Carson, D. K. (2000). Improving the positive predictive value of screening for developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *43*(4), 821-833.

Leonard, L. B. (2014). *Children with specific language impairment*. MIT press.

Levine, D., Pace, A., Luo, R., Hirsh-Pasek, K., Golinkoff, R. M., de Villiers, J., ... & Wilson, M. S. (2020). Evaluating socioeconomic gaps in preschoolers' vocabulary, syntax and language process skills with the Quick Interactive Language Screener (QUILS). *Early Childhood Research Quarterly*, *50*, 114-128.

McCauley, R. J., & Swisher, L. (1984). Psychometric review of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders*, *49*(1), 34-42.

Moland, C. W. (2011). Comparison of the Diagnostic Evaluation of Language Variation-Screening Test (DELV-ST) to two other screeners for low-income, African American children. Louisiana State University, Baton Rouge, LA. Retrieved from https://digitalcommons.lsu.edu/gradschool_dissertations/2150/

National Center for Education Statistics (NCES) Home Page, *U.S. Department of Education*. (n.d.). Retrieved from http://nces.ed.gov/

No Child Left Behind Act of 2006, P.L. 107-110, 20 U.S.C. (2006).

Oetting, J. B., Gregory, K. D., & Rivière, A. M. (2016). Changing how speech-language pathologists think and talk about dialect variation. *Perspectives of the ASHA Special Interest Groups*, *1*(16), 28-37.

Oetting, J. B., McDonald, J. L., Seidel, C. M., & Hegarty, M. (2016). Sentence recall by children with SLI across two nonmainstream dialects of English. *Journal of Speech, Language, and Hearing Research*, *59*(1), 183-194.

Plante, E., & Vance, R. (1995). Diagnostic accuracy of two tests of preschool language. *American Journal of Speech-Language Pathology*, *4*(2), 70-76.

Pruitt, S., & Oetting, J. (2009). Past tense marking by African American English–speaking children reared in poverty. *Journal of Speech, Language, and Hearing Research*.

Qi, C. H., Kaiser, A. P., Milan, S., & Hancock, T. (2006). Language performance of low-income African American and European American preschool children on the PPVT–III. *Language, Speech, and Hearing Services in Schools*.

Qi, C. H., Kaiser, A. P., Milan, S. E., Yzquierdo, Z., & Hancock, T. B. (2003). The performance of low-income, African American children on the Preschool Language Scale—3. *Journal of Speech, Language, and Hearing Research*.

Redmond, S. M., Ash, A. C., Christopulos, T. T., & Pfaff, T. (2019). Diagnostic Accuracy of Sentence Recall and Past Tense Measures for Identifying Children's Language Impairments. *Journal of Speech, Language, and Hearing Research*, *62*(7), 2438-2454.

Rice, M. L. (2020). Clinical lessons from studies of children with specific language impairment. *Perspectives of the ASHA Special Interest Groups*, *5*(1), 12-29.

Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *37*(1), 106-122.

Rice, M. L., Taylor, C. L., Zubrick, S. R., Hoffman, L., & Earnest, K. K. (2020). Heritability of specific language impairment and nonspecific language impairment at ages 4 and 6 years across phenotypes of speech, language, and nonverbal cognition. *Journal of Speech, Language, and Hearing Research*, 1-21.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(6), 1412-1431.

Sittner Bridges, M., & Catts, H. W. (2011). The Use of a Dynamic Screening of Phonological Awareness to Predict Risk for Reading Disabilities in Kindergarten Children. Journal of Learning Disabilities, 44(4), 330–338. https://doi.org/10.1177/0022219411407863

Snowling, M. J., Bishop, D. V. M., Stothard, S. E., Chipchase, B., & Kaplan, C. (2006). Psychosocial outcomes at 15 years of children with a preschool history of speech-language impairment. *Journal of Child Psychology and Psychiatry*, *47*(8), 759-765.

Spaulding, T. J., Plante, E., & Farinella, K. (2006). Eligibility for language impairment: Is the low end of normal always appropriate? Language, Speech, and Hearing Services in the School, 37, 61-72.

Taylor, C. L., Christensen, D., Lawrence, D., Mitrou, F., & Zubrick, S. R. (2013). Risk factors for children's receptive vocabulary development from four to eight years in the Longitudinal Study of Australian Children. *PLOS one*, *8*(9).

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of speech, language, and hearing research*, *40*(6), 1245-1260.

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1284-1294.

Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. *Developmental science*, *11*(1), 84-97.

Weiler, B., Schuele, C. M., Feldman, J. I., & Krimm, H. (2018). A multiyear population-based study of kindergarten language screening failure rates using the Rice Wexler Test of Early Grammatical Impairment. *Language, speech, and hearing services in schools*, *49*(2), 248-259.

Whitehouse, A. J., Watt, H. J., Line, E. A., & Bishop, D. V. (2009). Adult psychosocial outcomes of children with specific language impairment, pragmatic language impairment and autism. *International journal of language & communication disorders*, *44*(4), 511–528. https://doi.org/10.1080/13682820802708098

Zhang, X., & Tomblin, J. B. (2000). The association of intervention receipt with speech-language profiles and social-demographic variables. *American Journal of Speech-Language Pathology*, *9*(4), 345-357.

APPENDIX

LANGUAGE QUESTIONNAIRE

to accompany the QUICK INTERACTIVE LANGUAGE SCREENER™ (QUILS™)

Completed by: _____ Date: _____

Student's name: _____

Relationship to student: _____

This questionnaire should be used when a student hears or speaks more than one language at home. For each question, please check the one box that best describes your family.

	Only English	Mostly English	English and another language	Mostly other language	Only other language
What language does the primary caregiver use when speaking to the child?	☐	☐	☐	☐	☐
What language does the secondary caregiver use when speaking to the child?	☐	☐	☐	☐	☐
If there are siblings, what language do they use when speaking to the child?	☐	☐	☐	☐	☐
What language does the child use when talking to the primary caregiver at home?	☐	☐	☐	☐	☐
What language does the child use when talking to the secondary caregiver at home?	☐	☐	☐	☐	☐
What language does the child use when talking to the siblings at home?	☐	☐	☐	☐	☐
What language does the child use when talking to friends outside the home?	☐	☐	☐	☐	☐

LANGUAGE QUESTIONNAIRE

to accompany the
QUICK INTERACTIVE LANGUAGE SCREENER™ (QUILS™)

Completed by:_____ Date: _____

Student's name: _____

Relationship to student: _____

This questionnaire should be used when a student hears or speaks more than one language at home. For each question, please check the one box that best describes your family.

¿Qué idioma usa cuando habla con el niño la persona que lo cuida principalmente?

Solo Inglés	Mayoría Inglés	Inglés y otro idioma	Mayoría otro idioma	Solo otro idioma
☐	☐	☐	☐	☐

¿Qué idioma usa cuando habla con el niño la persona que lo cuida de manera secundaria?

Solo Inglés	Mayoría Inglés	Inglés y otro idioma	Mayoría otro idioma	Solo otro idioma
☐	☐	☐	☐	☐

Si hay hermanos, ¿qué idioma usan cuando hablan con el niño?

Solo Inglés	Mayoría Inglés	Inglés y otro idioma	Mayoría otro idioma	Solo otro idioma
☐	☐	☐	☐	☐

¿Qué idioma usa el niño cuando habla con la persona que lo cuida en casa principalmente?

Solo Inglés	Mayoría Inglés	Inglés y otro idioma	Mayoría otro idioma	Solo otro idioma
☐	☐	☐	☐	☐

¿Qué idioma usa el niño cuando habla con la persona que lo cuida de manera secundaria en casa?

Solo Inglés	Mayoría Inglés	Inglés y otro idioma	Mayoría otro idioma	Solo otro idioma
☐	☐	☐	☐	☐

¿Qué idioma usa el niño cuando habla con sus hermanos en casa?

Solo Inglés	Mayoría Inglés	Inglés y otro idioma	Mayoría otro idioma	Solo otro idioma
☐	☐	☐	☐	☐

¿Qué idioma usa el niño cuando se comunica con amigos fuera de casa?

Solo Inglés	Mayoría Inglés	Inglés y otro idioma	Mayoría otro idioma	Solo otro idioma
☐	☐	☐	☐	☐